Extremophiles (2015) 19:395–406
DOI 10.1007/s00792-014-0725-y

ORIGINAL PAPER

Insights into the interaction between Cren7 and DNA: the role of loop β3–β4

Zhenfeng Zhang · Yong Gong · Yuanyuan Chen · Hongbin Li · Li Huang

Received: 27 September 2014 / Accepted: 11 December 2014 / Published online: 3 January 2015

Abstract *Sulfolobus* synthesizes large amounts of small chromatin proteins Cren7 and Sul7d. The two proteins share overall structural similarity, but differ distinctly in the DNA-binding region between β3- and β4-strands. While Sul7d possesses a hinge of two amino acid residues, Cren7 contains a flexible seven-residue loop (loop β3–β4) in the region. Here, we report the role of loop β3–β4 in the interaction of Cren7 with duplex DNA. We show that all residues with a large side chain on the loop, i.e., Pro30, Lys31, Arg33 and Lys34, contributed significantly to the binding of Cren7 to DNA. The three basic amino acids affected the ability of Cren7 to constrain negative DNA supercoils in a residue number-dependent manner. The crystal structure of a complex between a mutant Cren7 protein (GR) with loop β3–β4 replaced by two residues (Gly and Arg) to mimic the hinge at the corresponding position in Sul7d and an 8-bp dsDNA has been determined. Structural comparison between the GR–DNA and Cren7–DNA complexes shows that GR resembles Sul7d more than Cren7 in DNA-binding size and in the effect on the width of the major groove of DNA and the pattern of DNA bending. However, GR induces smaller DNA curvature than Sul7d. Our results suggest that Cren7 and Sul7d package chromosomal DNA in a slightly different fashion, presumably permitting different chromosomal accessibility by proteins functioning in DNA transactions.

Keywords Chromatin protein · Cren7 · Loop β3–β4 · Protein–DNA complex · *Sulfolobus solfataricus* · Archaea

The atomic coordinates and structure factors for the GR–GTGATCAC and the Cren7–GTGATCAC complexes have been deposited in the Protein Data Bank (http://www.pdb.org) under accession codes 4R55 and 4R56, respectively.

Communicated by H. Atomi.

Electronic supplementary material The online version of this article (doi:10.1007/s00792-014-0725-y) contains supplementary material, which is available to authorized users.

Z. Zhang · H. Li · L. Huang (✉)
State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, No. 1 West Beichen Road, Chaoyang District, Beijing 100101, China
e-mail: huangl@sun.im.ac.cn

Y. Gong
Center for Multi-disciplinary Research, Institute of High Energy Physics, Chinese Academy of Sciences, 19B YuquanLu, Shijingshan District, Beijing 100049, China

Y. Chen
National Key Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Chaoyang District, Beijing 100101, China

Introduction

Genomic DNA must be compacted to be accommodated by the cell while kept accessible to molecular machineries for DNA transactions. Chromatin proteins are responsible for the compaction and organization of the chromosomal DNA (Luijsterburg et al. 2008). Eukarya encode histones, which wrap DNA around the core octamer to form a nucleosome (Luger et al. 1997). In bacteria, the chromosomal DNA is compacted into a nucleoid by a number of small DNA-binding proteins, such as HU, IHF, H-NS, Fis and Lrp (Dame 2005; Wu 2004). However, the situation is more complex in Archaea as different archaeal phyla may employ different sets of DNA-binding proteins (Zhang et al. 2010). Alba is a unique nucleic acid-binding protein highly conserved among the

domain Archaea (Bell et al. 2002). However, the protein appears to bind both DNA and RNA (Guo et al. 2003; Marsh et al. 2005). Therefore, the physiological roles of Alba remain to be understood. Euryarchaea produce archaeal histones capable of packaging DNA in a similar manner to their eukaryal counterparts (Reeve 2003; Sandman and Reeve 2005). However, archaeal histones are absent in crenarchaea with the exception of some species (Cubonova et al. 2005). Several small DNA-binding proteins, including Sul7d, CC1 and Cren7, have been isolated from crenarchaea (Grote et al. 1986; Guo et al. 2008; Luo et al. 2007). While Sul7d and CC1 are phylogenetically restricted to some crenarchaeal species, Cren7 is highly conserved in the phylum (Zhang et al. 2012). It is suggested that crenarchaea may have evolved a Cren7-based DNA packaging strategy.

Cren7, a small and basic chromatin protein, preferentially binds dsDNA over ssDNA, protects DNA duplex from thermal denaturation and constrains negative DNA supercoils in vitro (Guo et al. 2008). Sul7d, a 7-kDa chromatin protein from *Sulfolobus*, resembles Cren7 in biochemical properties (Mai et al. 1998; Napoli et al. 2002). Both Cren7 and Sul7d compact DNA by inducing rigid bends with similar efficiencies (Driessen et al. 2013). In *S. solfataricus* P2 cells, Cren7 and Sso7d, a member of the Sul7d family, are both synthesized in abundance, amounting to ~1 and ~5 %, respectively, of the total cellular proteins (Guo et al. 2008). Therefore, it appears that the two proteins are both functional and possibly serve non-redundant roles in vivo.

Crystal structures of Cren7 and Sul7d alone and in complex with duplex DNA have been determined (Gao et al. 1998; Robinson et al. 1998; Zhang et al. 2010). Despite the low similarity at the amino acid sequence level, the two proteins share a similar tertiary structure. However, comparison of the DNA-binding surfaces of Cren7 and Sac7d, a member of the Sul7d family from *S. acidocaldarius*, reveals a remarkable difference in the loop region between β3 and β4 strands (loop β3–β4) (Zhang et al. 2010). A structure-based sequence alignment shows that Cren7 possesses a long (seven residues) and highly flexible loop β3–β4, whereas Sac7d contains a small hinge (two residues) at the corresponding position (Fig. 1). Loop β3–β4 of Cren7 undergoes a significant conformational change upon binding of the protein to DNA and covers about two base pairs in the crystals of the Cren7–DNA complex. The loop region is conserved among Cren7 homologs, and most of the residues on the loop interact directly with the DNA strands.

To elucidate the role of loop β3–β4 in chromosomal organization by Cren7 and learn more about the potential functional differences between Cren7 and Sul7d, it is essential to understand the effect of the loop on the interaction of Cren7 with DNA. In the present study, we introduced point mutations and deletions into the loop region and determined the effects of the mutations on DNA binding and supercoil constraining by Cren7. In addition, we solved the crystal structure of a complex between a Cren7 deletion variant (GR), which mimics Sac7d in the loop region, with a duplex DNA fragment. Our results suggest that Cren7 and Sul7d may allow chromosomal DNA to be

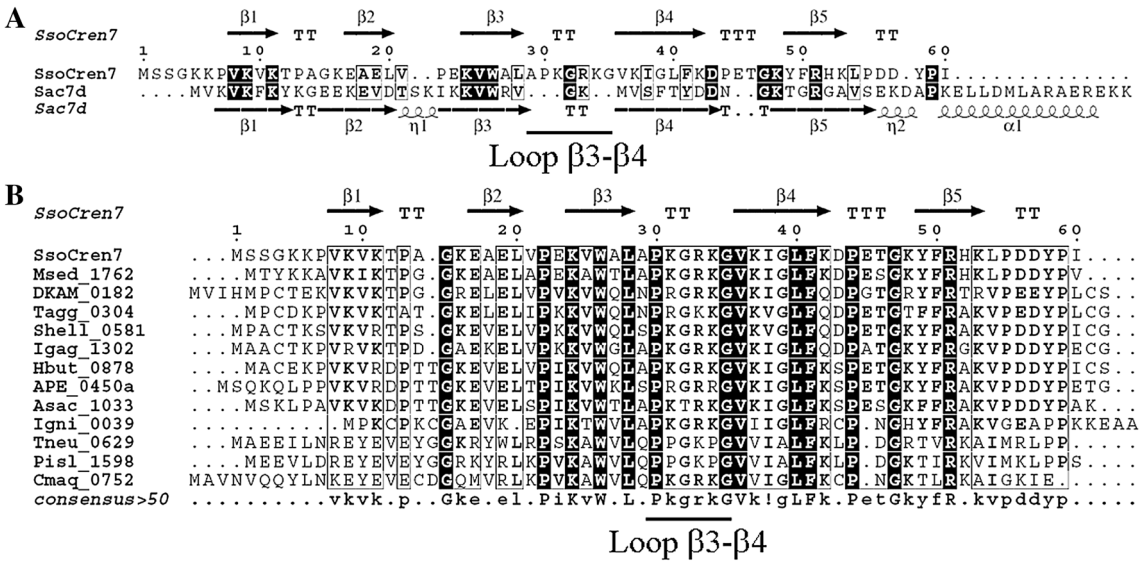

Fig. 1 Sequence alignment of Cren7 homologs and Sac7d. **a** Structure-based sequence alignment of Cren7 (PDB code: 3LWH) and Sac7d (PDB code: 1AZQ). The secondary structures of both proteins are indicated. **b** Sequence alignment of selected Cren7 homologs from representative crenarchaeal genera. The secondary structures of Cren7 are indicated. The loop β3–β4 regions of Cren7 in both panels are marked

packaged in a slightly different fashion and, therefore, to exhibit different accessibility.

Materials and methods

Preparation of mutant proteins

Point mutations were introduced into Cren7 by site-directed mutagenesis using the Fast Mutagenesis System (TransGen Biotech, China) with pET30a–Cren7, an expression plasmid for Cren7 (Guo et al. 2008), as the template. The deletion mutants of Cren7 were prepared by overlapping PCR and subsequently inserting the resulting PCR fragment into pET30a (Novagen, USA). Primers used in the above assays are listed in supplementary table S1. All the mutations were confirmed by DNA sequencing. The mutant constructs were transformed into *E. coli* BL21 (DE3) cells, and recombinant proteins were overproduced and purified as described previously (Zhang et al. 2010). The purified proteins were lyophilized, resuspended in 20 mM Tris–Cl, pH 6.8, and 1 mM EDTA, and stored at −80 °C.

Surface plasmon resonance assays (SPR)

SPR experiments were carried out at 25 °C using the BIAcore 3000 instrument (BIAcore AB, Uppsala, Sweden). The running buffer contained 50 mM Tris–Cl, pH 7.5, 100 mM NaCl and 0.005 % (v/v) Tween 20. Two complementary oligonucleotides (supplementary table S1), including one with a biotin label, were immobilized on the SA sensor chip to 91–97 response units (RU). A blank flow cell was used as the reference to correct for instrumental and concentration effects. Wild-type or mutant Cren7 at a concentration within a range spanning the K_D value for the interaction of the protein with the DNA was injected over the DNA surface for 2 min at a flow rate of 30 µl/min. After the dissociation phase (2–4 min), bound protein was removed by a 30-s wash with 0.01 % SDS, followed by a 60-s buffer injection. Measurement at the lowest protein concentration in each SPR experiment was repeated once. Equilibrium and kinetic constants were calculated by a global fit to 1:1 Langmuir binding model (BIA evaluation 4.1 software).

Nick-closure assays

Nick-closure assays were carried out as described previously with some modifications (Guo et al. 2008). Plasmid pBR322 containing a single nick per circular molecule was prepared by digestion with the nicking enzyme Nb.Bpu10I (Fermentas). The nicked plasmid (0.2 µg) was mixed with various amounts of wild-type or mutant Cren7 in 40 mM Tris–Cl, pH 7.8, 10 mM $MgCl_2$, 10 mM DTT, 0.5 mM ATP and 100 µg/ml BSA, and incubated at 25 °C for 10 min. Ligation was then initiated by the addition of T4 DNA ligase (2U; Fermentas) to the mixture to a final volume of 11 µl. After further incubation at 25 °C for 2 min, reactions were terminated with 5× stop buffer (250 mM EDTA and 2.5 % SDS). Samples were deproteinized and analyzed by agarose gel electrophoresis in the presence or absence of chloroquine. Gels were stained in 0.5 µg/ml ethidium bromide and visualized under UV light.

Size exclusion chromatography

A 100-µl sample (~1 mM) of wild-type or mutant Cren7 was subjected to size exclusion chromatography in 0.1 M NH_4Ac buffer (pH 5.2) on a Superdex 75 column (GE healthcare) at a flow rate of 0.5 ml/min at 4 °C. Elution of the protein was monitored by absorbance at 280 nm. The void volume (v_o) and total volume (v_t) of the column were determined with blue dextran and xylene cyanol, respectively. The fractional retention (K_{av}) was calculated using the formula $K_{av} = (v_e - v_o)/(v_t - v_o)$, where v_e is the peak elution volume. A standard curve of K_{av} against log M_r was generated by using low molecular weight protein standards (GE healthcare).

Crystallization, data collection and structure determination

Cocrystallization screening was performed with GR, a deletion variant of Cren7, and each of the three 8-bp palindromic DNA sequences (supplementary table S1) by sitting-drop vapor diffusion. Crystals suitable for the X-ray diffraction were obtained only with the GR–$d(GTGATCAC)_2$ complex. GR and a DNA duplex were mixed at a molar ratio of 2:1 to a final concentration of 2 mM for the protein. The sample (1 µl) was mixed with the reservoir solution (1 µl) and equilibrated at 20 °C with 0.2 M NH_4Ac, 0.01 M $MgAc_2$, 0.05 M sodium cacodylate trihydrate, pH 6.5, and 30 % PEG8000. Crystallization of Cren7–$d(GTGATCAC)_2$ was performed under the conditions identical to those described previously for the Cren7–DNA complexes (Zhang et al. 2010). The crystals were mounted on nylon loops and immediately frozen in liquid nitrogen. The data for the complex was collected to 1.8 Å at the Rigaku MicroMax-007HF (Rigaku, Japan) at 100 K, and the wavelength was 1.5418 Å. The data were integrated and scaled with HKL2000 (Otwinowski and Minor 1997).

The structures of Cren7–$d(GTGATCAC)_2$ and GR–$d(GTGATCAC)_2$ were determined by molecular replacement using the program Phaser (McCoy et al. 2007) from the CCP4 program suite (Collaborative Computational Project N 1994) using the crystal structure of Cren7–$d(GTGATCAC)_2$ (PDB code: 3LWH) as the initial model. The software packages Refmac (Murshudov et al. 1997)

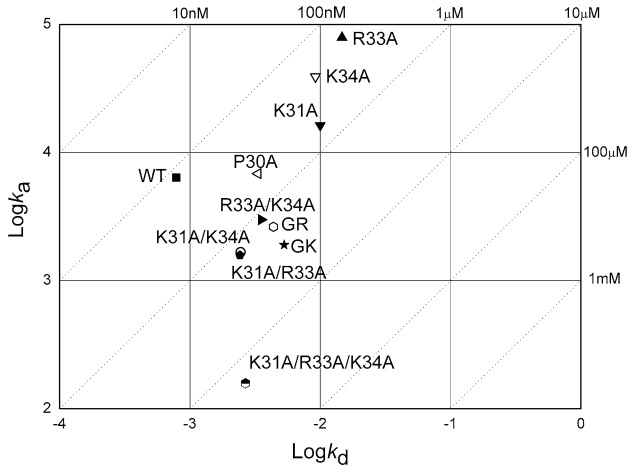

Fig. 2 Kinetic characterization of the binding of wild-type and mutant Cren7 to DNA by SPR. Wild-type or mutant Cren7 was injected over immobilized DNA. The k_a (recognition) is plotted against the k_d (complex stability) on a logarithmic scale. The *diagonal lines* (*gray dotted lines*) are isoaffinity lines representing the affinity $K_D = k_d/k_a$

Table 1 Amino acid sequences of loop β3–β4 and its mutant forms

Protein	Sequence
Cren7	28-L APKGRKG V-36
LD	28-L A----------G V-29
GR	28-L ------GR---- V-31
GK	28-L ------GK---- V-31
Sac7d	26-V GK M-29
P30A	28-L AAKGRKG V-36
K31A	28-L APAGRKG V-36
R33A	28-L APKGAKG V-36
K34A	28-L APKGRAG V-36
K31A/R33A	28-L APAGAKG V-36
K31A/K34A	28-L APAGRAG V-36
R33A/K34A	28-L APKGAAG V-36
K31A/R33A/K34A	28-L APAGAAG V-36
G32A	28-L APKARKG V-36
G35A	28-L APKGRKA V-36
ΔG32	28-L APK--RKG V-36
ΔG35	28-L APKGRK-- V-36

and Coot (Emsley and Cowtan 2004) were used to complete the model. Statistics for data collection and refinement are given in Table 3. The qualities of the final structures were evaluated using PROCHECK (Laskowski et al. 1993). DNA conformations were analyzed using the program X3DNA (Lu and Olson 2008). All images were prepared using Pymol (http://www.pymol.org) and ESPript (Gouet et al. 1999).

Results

Kinetic analysis of the interaction of loop β3–β4 mutants with DNA

Loop β3–β4 of Cren7 comprises seven amino acid residues, i.e., Ala29, Pro30, Lys31, Gly32, Arg33, Lys34 and Gly35. We have previously shown that single-alanine substitution of Pro30, Lys31 or Arg33, each of which is located on the DNA-binding surface of Cren7, resulted in significant reduction in binding affinity of the protein for dsDNA (Zhang et al. 2010). To learn more about the role of loop β3–β4 in the interaction of Cren7 with DNA, we first introduced a single-alanine substitution at Pro30, Lys31, Arg33 or Lys34 and double- and triple-alanine substitutions at the latter three basic amino acid residues, and compared the ability of these mutants to bind DNA and constrain negative DNA supercoils (supplementary figure S1, Fig. 2; Tables 1, 2). All the mutants showed reduced binding affinity for DNA, as compared to that of wild-type Cren7. The four single mutants, i.e., P30A, K31A, R33A and K34A, were similar in binding affinity with K_D values in the 10^{-7} M range. However, the scattered distribution of the mutants in the k_a/k_d map suggests that these amino acid residues may differ in their roles in DNA binding by Cren7. The association rate constant (k_a) of P30A was nearly identical to that of wild-type Cren7, but the dissociation rate constant (k_d) of the mutant was fourfold as high as that of the wild-type protein. Presumably, the hydrophobic interactions between the side chain of Pro30 and the bases of the DNA serve to stabilize the Cren7–DNA complex. On the other hand, mutation of each of the three basic amino acid residues was accompanied by a 10- to 20-fold increase in both k_a and k_d and, therefore, only a slight decrease in DNA-binding affinity. Among the three single mutations, substitution of Ala31 for Lys31 resulted in the largest decrease (~5 fold) in DNA-binding affinity, pointing to the importance of Lys31 in the interaction of the protein with the DNA. R33A showed a slight reduction (~1.3 fold) in DNA binding, because its rates of both association and dissociation increased to similar extents. Notably, mutation of Lys34, which is not in contact with the DNA in the crystals (Zhang et al. 2010), also led to a twofold decrease in the binding affinity of the protein, raising the possibility that the residue interacts with the DNA in solution. This is consistent with the observation that alanine substitution of both Lys34 and Lys31 or Arg33 further reduced the binding affinity by two- or fivefold, respectively, as compared to that of K31A or R33A. In fact, the K_D values of the three double mutants (K31A/R33A, K31A/K34A and R33A/K34A) were similar (1.2–1.5 μM). Moreover, the double mutants were clustered in the k_a/k_d map, exhibiting a tenfold decrease

Table 2 Kinetic analysis of DNA binding by wild-type and mutant Cren7 by SPR

Mutant	Affinity K_D (M)	Association rate k_a ($M^{-1}s^{-1}$)	Dissociation rate k_d (s^{-1})
Cren7–WT	1.34×10^{-7}	$6.13 \pm 0.14 \times 10^{3}$	$8.21 \pm 0.34 \times 10^{-4}$
LD[a]	2.10×10^{-5}	ND[b]	ND[b]
GR	1.67×10^{-6}	$2.63 \pm 0.04 \times 10^{3}$	$4.38 \pm 0.06 \times 10^{-3}$
GK	2.78×10^{-6}	$1.90 \pm 0.06 \times 10^{3}$	$5.29 \pm 0.08 \times 10^{-3}$
P30A	4.73×10^{-7}	$6.59 \pm 0.10 \times 10^{3}$	$3.12 \pm 0.04 \times 10^{-3}$
K31A	6.46×10^{-7}	$1.57 \pm 0.04 \times 10^{4}$	$1.01 \pm 0.02 \times 10^{-2}$
R33A	1.79×10^{-7}	$7.75 \pm 0.16 \times 10^{4}$	$1.39 \pm 0.02 \times 10^{-2}$
K34A	2.36×10^{-7}	$3.91 \pm 0.03 \times 10^{4}$	$9.20 \pm 0.06 \times 10^{-3}$
K31A/R33A	1.53×10^{-6}	$1.58 \pm 0.03 \times 10^{3}$	$2.42 \pm 0.05 \times 10^{-3}$
K31A/K34A	1.46×10^{-6}	$1.67 \pm 0.06 \times 10^{3}$	$2.45 \pm 0.16 \times 10^{-3}$
R33A/K34A	1.18×10^{-6}	$2.99 \pm 0.09 \times 10^{3}$	$3.53 \pm 0.11 \times 10^{-3}$
K31A/R33A/K34A	1.69×10^{-5}	$1.58 \pm 0.05 \times 10^{2}$	$2.67 \pm 0.06 \times 10^{-3}$

[a] Equilibrium and kinetic rate constants for this mutant were calculated using a steady-state binding model (BIA evaluation 4.1 software)

[b] Not determined

in DNA-binding affinity, as a result of a slight decrease in both k_a and k_d values, as compared to those of the wild-type protein. Substitution of all three basic residues with alanine reduced the binding affinity of Cren7 for DNA by over two orders of magnitude, emphasizing the critical role of the positively charged residues on loop β3–β4 in the interaction of the protein with the DNA. Intriguingly, the drastic loss in DNA binding by K31A/R33A/K34A was due primarily to a 40-fold decrease in the k_a of the protein. A threefold reduction was observed in the dissociation rate constant of the triple mutant. It is worth noting that both the double and the triple mutants had similar dissociation rate constants. These results suggest that the positive surface electric potential in the loop β3–β4 region plays a key role in the Cren7–DNA interaction and contributes more to the association of Cren7 to DNA than to the stabilization of the Cren7–DNA complex.

Contribution of the basic residues in loop β3–β4 to constraining of negative DNA supercoils by Cren7

Cren7 is highly efficient in constraining DNA into negative supercoils (Guo et al. 2008). Nick-closure assays were performed to determine the ability of the above loop β3–β4 mutants to constrain DNA supercoils (Fig. 3a). Singly nicked pBR322 was ligated in the absence or presence of wild-type or mutant Cren7. It was noticed that the plasmid nick closed by T4 DNA ligase at 25 °C appeared to be slightly positively supercoiled rather than relaxed. This presumably resulted from the small difference in the experimental conditions (e.g., temperature, ionic strength) of the nick ligation assay and electrophoresis in agarose gel (Mai et al. 1998). As increasing amounts of the protein were added to the reaction and thus increasing amounts of negative supercoils were constrained by the protein, the ligated plasmid first lost positive supercoils and subsequently accumulated negative supercoils. The linking changes of the ligated pBR322 were measured by band counting and plotted against the protein/DNA mass ratios (Fig. 3b). The slope of each plot is taken as a measure of the efficiency of the protein in constraining DNA supercoils.

As shown in Fig. 3b, ~23 negative supercoils were constrained by Cren7 at the mass ratio of 1:1 (molecular ratio of ~436:1). This translates into binding of 207 to 253 Cren7 molecules to a plasmid molecule according to the stoichiometry of 9–11 Cren7 molecules for a constrained supercoil (Zhang et al. 2010). Given a DNA-binding site size of ~8 bp, ~545 Cren7 molecules are required to saturate the 4361-bp pBR322 DNA molecule. Therefore, the binding of Cren7 to the plasmid DNA was far from saturation at the mass ratio of 1:1. This is consistent with our previous finding that Cren7 did not saturate the pBR322 DNA until the mass ratio reached 8:1 (Guo et al. 2008). P30A constrained DNA supercoils as efficiently as wild-type Cren7. Since the former was fourfold weaker than the latter in binding to DNA, it seems that the ability of Cren7 to constrain DNA supercoils is not directly related to the affinity of the protein to DNA. All of the remaining alanine substitution mutants of Cren7 showed significantly reduced ability to constrain negative DNA supercoils, as compared to that of the wild-type protein. There appears to be a general correlation between the number of basic residues on loop β3–β4 and the supercoil-constraining ability of the protein variants. Cren7 containing one, two or three alanine substitutions at Lys31, Arg33 and Lys34 residues was approximately 2-, 4- or 16-fold, respectively, less efficient in constraining DNA supercoils than the wild-type protein. In other words, the efficiency of supercoil constraining by the protein decreased by approximately fourfold with the loss of each basic residue on loop β3–β4. Notably, little difference in DNA supercoiling was observed between the three single mutants or between the two double mutants, although each of the two sets of the mutants differed among themselves in DNA-binding affinity. Therefore, we infer

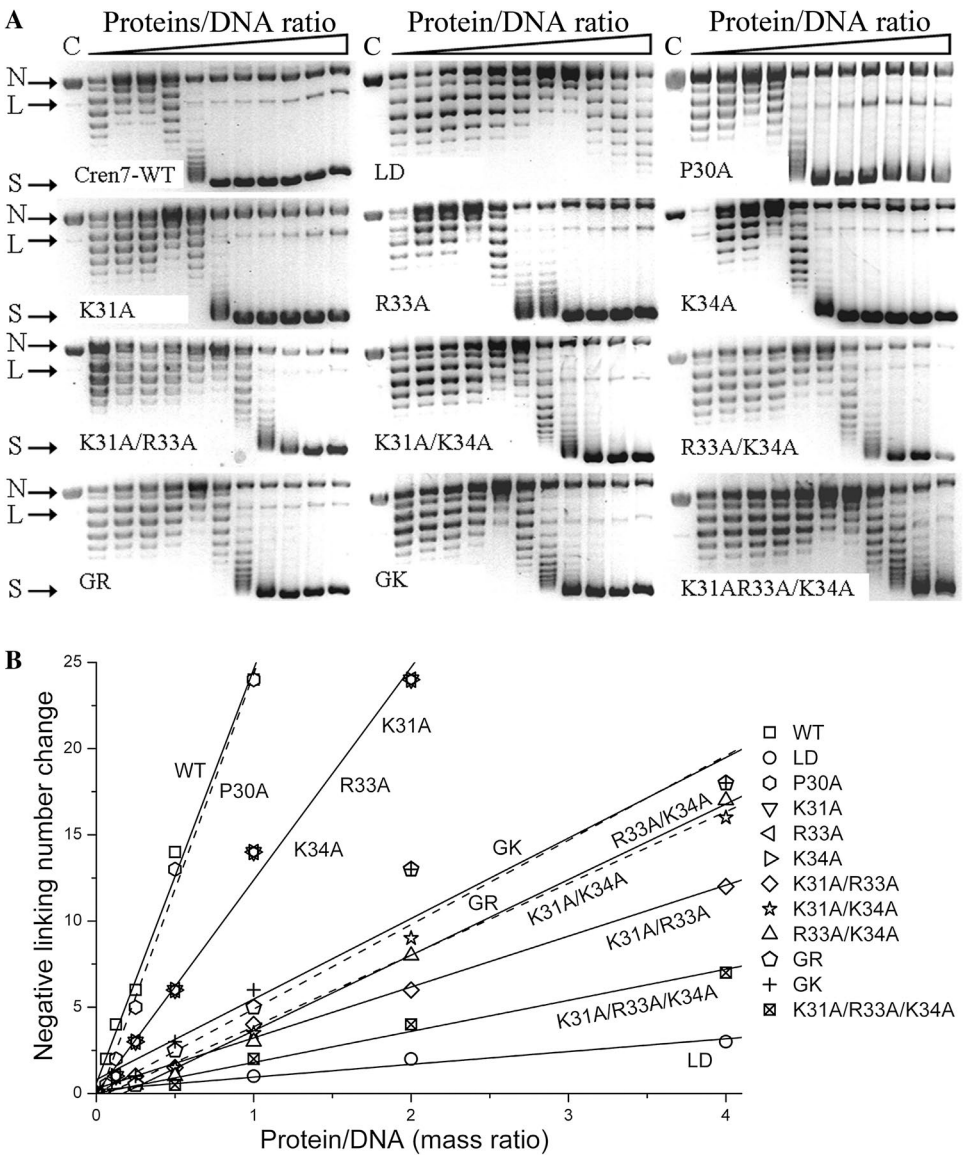

Fig. 3 DNA nick-closure assays. **a** Effect of mutations in loop β3–β4 on the ability of Cren7 to constrain negative DNA supercoils. Plasmid DNA containing a single nick was incubated in the presence of increasing amounts of wild-type or mutant Cren7 before ligation by T4 DNA ligase. Protein/DNA mass ratios were 0, 0.0625, 0.125, 0.25, 0.5, 1, 2, 4, 8, 16 and 32, respectively. *C* pBR322 with a single nick, *N* nicked pBR322 DNA, *L* linear pBR322 DNA, *S* negatively supercoiled pBR322 DNA. **b** A plot of the linking number change against the protein/DNA mass ratio by linear fit with the software Origin 7.5. The plots for K31A, R33A and K34A are overlapping and thus shown as a *single solid line*; those for P30A, GR and K31A/K34A are shown as a *dashed line*; and those for the rest of the proteins are shown as a *solid line*. The linking number change of pBR322 nick closed in the presence or absence of a Cren7 mutant was measured by resolving plasmid topoisomers on an agarose gel, as shown in Fig. 2a, and subsequent band counting

from these observations that the positively charged residues in the loop region exert a similar impact on DNA geometry.

The role of the glycine residues on loop β3–β4 in the folding of Cren7

Among the seven residues in loop β3–β4, two glycine residues (Gly32 and Gly35) are highly conserved in the Cren7 homologs (Fig. 1), suggesting their importance to the structure and/or function of Cren7. To understand the potential roles of the two residues, we performed alanine substitution and deletion analyses on each of the two glycine residues (Table 1). Four mutants, i.e., G32A, G35A, ΔG32 and ΔG35, resembled wild-type Cren7 in both DNA binding and supercoiling (Data not shown). This is not unexpected since the two glycine residues are not in contact with DNA in the Cren7–DNA cocrystals (Zhang et al. 2010). Surprisingly, however, an unexpected protein peak, in addition to the typical monomeric Cren7 peak, was found when G32A, ΔG32 and ΔG35 were subjected to size exclusion chromatography (Fig. 4). Since the protein variants eluted as monomers in the expected peak showed calculated molecular masses ranging from 8.86 to 9.05 kDa, the protein species that came off the column in the additional peak appeared to be dimers based on their molecular masses (18.6–21.0 kDa in size). Substitution of Gly32 with alanine facilitated dimerization of the mutant, with ~15 % of the G32A molecules existing as dimers under our experimental conditions. Deletion of Gly32 increased the formation of dimers to ~30 %. G35A existed as monomers, while about half of the ΔG35 molecules formed dimers. Therefore, the two glycine residues may serve to promote proper folding of Cren7, allowing the protein to exist as a monomer.

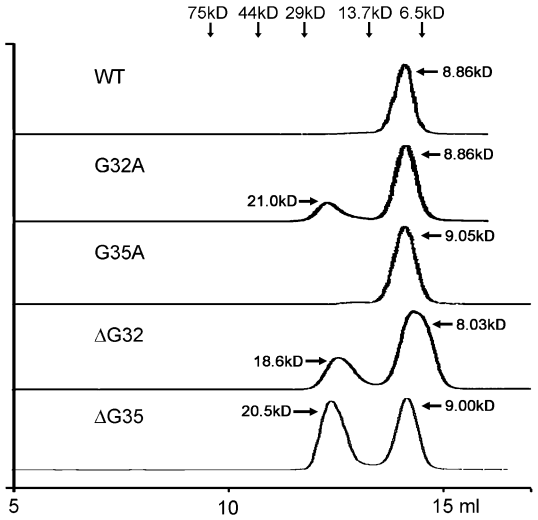

Fig. 4 Gel filtration chromatography of wild-type and mutant Cren7. All the protein samples were loaded at the same amount (~600 μg protein in 100 μl), as described in "Materials and Methods". The Y-axis represents the absorbance at 280 nm. The scale for the Y-axis for each elution profile was adjusted to permit easy comparison of the elution times of the proteins. Molecular weight markers (conalbumin, 75kD; ovalbumin, 44kD; carbonic anhydrase, 29kD; ribonuclease, 13.7kD; aprotinin, 6.5kD) are indicated. The calculated molecular masses of the sample peaks are also shown

Significance of the length of loop β3–β4

Unlike Cren7, Sac7d possesses a small hinge between β3 and β4 strands. Sequence alignment of the two structurally homologous proteins reveals the similarity between the two residues (Gly32–Arg33) in the middle of loop β3–β4 of Cren7 and the hinge region of Sac7d (Gly27–Lys28). A question then arises as to whether the function of loop β3–β4 in Cren7 can be maintained by a much shorter hinge similar to that in Sac7d. To answer this question, we prepared three deletion mutant proteins of Cren7, denoted LD, GR and GK, respectively (Table 1). LD, constructed by deleting the Pro30–Lys34 region, had a short loop consisting of an alanine residue and a glycine residue. To mimic the hinge between β3 and β4 strands in Sac7d, GR and GK were made by substituting the seven-residue β3–β4 loop for a two-residue stretch (i.e., Gly–Arg for GR and Gly–Lys for GK). DNA binding by these mutants was analyzed by SPR (Table 2). In agreement with a previous report (Feng et al. 2010), the K_D of LD was ~20 μM, which was nearly 200-fold higher than that of wild-type Cren7 (~123 nM). The substantial loss in the binding affinity of LD for DNA was almost entirely attributable to an increase in k_d, which was too high to be determined from the dissociation curves, indicating that the LD–DNA complex was very unstable. By comparison, the affinities of GR and GK for DNA were lower than the wild-type protein by only ~10- and ~20-fold, respectively. A slight decrease in k_a and a modest increase in k_d were observed for both mutant proteins. The difference between LD and GR (or GK) in DNA-binding affinity indicates that the positive charge in the loop region is crucial to the binding of Cren7 to DNA. Since LD binds to DNA less tightly than the triple substitution mutant (K31A/R33A/K34A), in which the three basic residues on loop β3–β4 were replaced with alanine, the size of the loop appears to contribute to the stability of the Cren7–DNA complex as well. This is also supported by the resemblance of both GR and GK to the double-substitution mutants (K31A/R33A, K31A/K34A and R33A/K34A) in DNA-binding kinetics and affinity, as shown by the clustering of these mutants in the k_a/k_d map (Fig. 2).

The deletion variants of Cren7 were also assayed for their ability to constrain DNA supercoils. All three mutants were able to constrain negative DNA supercoils (Fig. 3). As expected, LD was 32-fold less efficient than the wild-type Cren7 in constraining DNA supercoils. By contrast, the ability of GR or GK to constrain DNA supercoils was about fourfold lower than that of Cren7. It should be noticed that the difference in the ability to constrain negative supercoils

Table 3 Data collection and refinement statistics

	GR–GTGATCAC	Cren7–GTGATCAC
Data collection		
Space group	$P2_12_12_1$	$C222_1$
Cell dimensions		
a, *b*, *c* (Å)	42.5, 41.8, 54.3	77.9, 77.9, 104.4
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	50–1.80 (1.86–1.80)	50–2.30 (2.38–2.30)
R_{merge}	0.05 (0.08)	0.04 (0.48)
$I/\sigma I$	40.6 (25.3)	37.6 (3.7)
Completeness (%)	99.5 (99.0)	97.9 (96.3)
Redundancy	6.7 (6.6)	7.2 (7.4)
Refinement		
Resolution (Å)	17–1.80	50–2.30
No. of reflections	8922	13431
R_{work}/R_{free}	0.219/0.252	0.200/0.232
No. of atoms		
Protein	423	926
DNA	322	644
Water	72	116
B-factors		
Protein	13.0	47.0
DNA	11.5	33.2
Water	15.6	42.1
Rms deviations		
Bond lengths (Å)	0.016	0.008
Bond angles (°)	1.7	1.3

between wild-type Cren7 and the three variants was slightly underestimated since the comparison was made at the same protein/DNA mass ratio and there was a small difference (~10 %) in molecular mass between wild-type Cren7 (~6.5 kDa) and the three variants (~6 kDa), which corresponded to the protein/DNA molar ratios of 436:1 for wild-type Cren7 and 472:1 for the deletion variants, at the mass ration of 1:1. In addition, the efficiencies of GK and GR in constraining DNA supercoils were similar despite their difference in DNA-binding affinity. Interestingly, both GR and GK constrained supercoils with slightly higher efficiencies than the double substitution mutants, although the former had lower affinity for DNA than the latter (Table 2). It is speculated that replacement of the long loop, which covers 2 bp of DNA (Zhang et al. 2010), with a short hinge in Cren7 would reduce the binding size of the protein, allow increased binding density and thus enhance the efficiency of the protein to constrain DNA supercoils. Taken together, our results show that replacement of loop β3–β4 with a short hinge lacking a positively charged amino acid residue significantly impaired the ability of Cren7 to bind DNA and to constrain DNA supercoils, and this ability of the protein was partially restored if the hinge region contained a positively charged residue.

Crystal structure of the GR–DNA complex

To determine the structural basis of the effect of replacement of loop β3–β4 with a short hinge on DNA binding and supercoil constraining by the protein, we attempted to co-crystallize GR with one of the three 8-bp double-stranded DNA fragments (supplementary table S1). The crystal structure of GR–d(GTGATCAC)$_2$ was solved and refined to 1.8 Å (Table 3; Fig. 5a). A single protein in

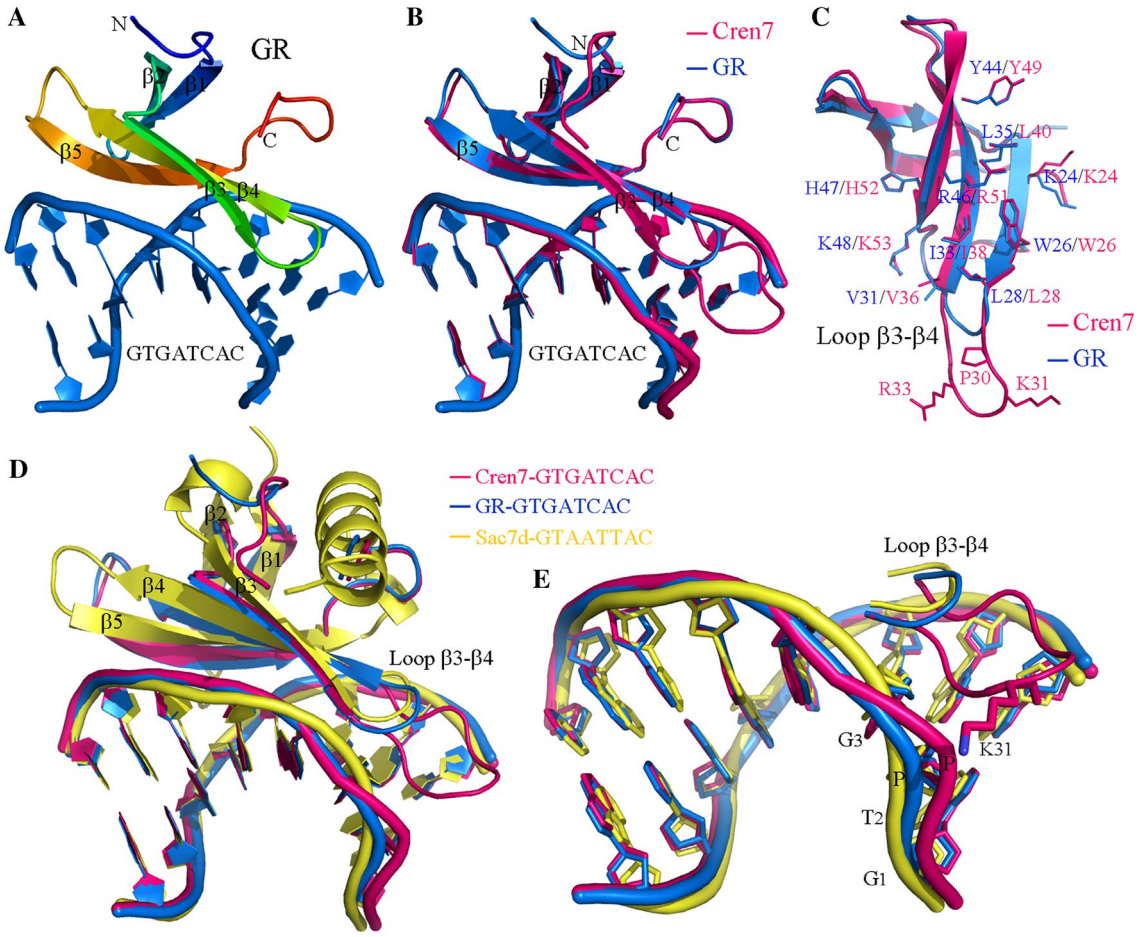

Fig. 5 Crystal structure of the GR–d(GTGATCAC)$_2$ complex. **a** Ribbon diagram of the GR–GTGATCAC complex. **b** Superposition of Cren7 and GR in complex with GTGATCAC, fitted by common protein main-chain atoms. **c** The DNA-binding surfaces of Cren7 and GR, with selected side chains labeled. The loop β3–β4 regions of the two proteins are highlighted, and the residues of Cren7 and GR are labeled in *pink and blue*, respectively. **d** Superposition of the Cren7–GTGATCAC (*pink*), GR–GTGATCAC (*blue*) and Sac7d–GTAATTAC (*yellow*; PDB code: 1AZQ) complexes fitted by all atoms from one of the DNA strands. **e** A detailed view of the DNA structures in **d**. The T2pA3 step is marked, and loop β3–β4 and Lys31 of Cren7 are also shown

complex with DNA is found in the asymmetric unit. The crystal structure of wild-type Cren7 in complex with the same DNA duplex was also resolved and refined to 2.3 Å (Table 3; Fig. 5a). Two Cren7–GTGATCAC complexes are in the asymmetric unit and the root-mean-square deviation (rmsd) value between them is 0.016 Å.

Superposition of GR with wild-type Cren7 in complex with DNA, fitted by their α-carbons with a root-mean-square distance of 0.326 Å, shows a similar overall structure with conformational differences in the regions of N-terminal five residues and loop β3–β4 (Fig. 5b). The GR–DNA complex also shares an overall structure with the Cren7–DNA complex. GR binds DNA as monomers in the minor groove. The bound protein induces a sharp kink (~48°) in the DNA duplex toward the major groove by intercalation of the side chain of Leu28 into the G3A4 step. The minor groove at the intercalating site is drastically widened to an A-form-like conformation. Remarkably, deletion of the long flexible loop leads to a reduction in binding size for GR (~6 bp), as compared to that for wild-type Cren7 (~8 bp) (supplementary figure S2). The short loop between β3 and β4 strands in GR remains fixed in the minor groove, but Arg30 on the loop does not interact with DNA, as evidenced by the poor electric density of the side chain of Arg30 in the crystal lattice. Structural comparison between GR and Sac7d in complex with DNA shows similar conformations in the region between β3 and β4 strands (Fig. 5d). All DNA-interacting residues are located on the triple-stranded β-sheet in GR. GR is similar to wild-type Cren7 in DNA-binding surface except for the conformational changes in Lys24 and Val31 side chains (Fig. 5c). Lys24 in GR contacts the phosphate group of C6 and the deoxyribose of T5 through hydrophobic interactions instead of forming hydrogen bond to the O1P of C6 through its NH_2 group, as found in the Cren7–DNA complex. Val31 of GR is still in close contact with the C14pA15 step of DNA, as in the Cren7–DNA complex. The side chain of Val31, however, is rotated by ~90° clockwise and its hydrophobic interaction with the deoxyribose of C14 no longer exists.

Deformation of DNA by bound GR

In the crystal structure, the DNA bound by GR is similar to that bound by wild-type Cren7 in global conformation except that the backbone in the T2pG3 region in the GR–DNA complex is not as deformed as in the wide-type protein–DNA complex (Fig. 5d, e). In fact, the local conformation of this region in the DNA bound by GR resembles that in the Sac7d–GTAATTAC complex. The phosphate group between T2 and G3 in the DNA bound by wild-type Cren7 is translocated as a result of the formation of a hydrogen bond to Lys31, resulting in distortion in the backbone. By comparison, the DNA backbone in the GR–DNA and Sac7d–DNA complexes is much smoother due to the lack of direct contacts between the loop and the backbone phosphate group.

As observed in the complex of wild-type Cren7 with DNA, binding of GR dramatically widens the minor groove to about 12.0 Å at the site of intercalation, changing the local DNA conformation from B-form (6.0 Å) to A-form (11.1 Å) (Arnott 1999). The major groove at this site is narrowed by about 2.6 Å to 9.0 Å, as compared to 11.6 Å in B-form DNA. In comparison, the width of the major groove at the intercalation site in the wild-type Cren7–GTGATCAC complex (12.0 Å) remains unchanged. Notably, GR alters minor and major grooves in the same manner as Sac7d, which widens the minor groove to ~12.9 Å and narrows the major groove to ~8.1 Å in the protein–GTAATTAC complex (Robinson et al. 1998).

A sharp kink occurs at the G3pA4 step in both the GR–DNA and the wild-type Cren7–DNA complexes. The kink angle of ~48° in the GR–DNA complex is even larger than that in the wild-type Cren7–DNA complex (~45°). However, the average roll angle of the DNA duplex bound by GR (~8.5°) is smaller than that bound by wild-type Cren7 (~9.9°), indicating that the bound DNA is overall less curved in the GR–DNA complex than in the wild-type protein–DNA complex (Table 4). The reduction in DNA curvature in the GR–DNA complex is mainly due to negative rolls at the G1pT2 and C6pA7 steps (about −4.0° and

Table 4 DNA helical parameters of the Cren7–GTGATCAC and GR–GTGATCAC complexes

Step	Cren7-GTGATCAC			GR-GTGATCAC		
	Rise (Å)	Roll (°)	Twist (°)	Rise (Å)	Roll (°)	Twist (°)
GT/AC	3.14	2.21	29.97	3.43	−4.03	30.14
TG/CA	2.93	8.56	27.61	2.83	7.94	30.45
GA/TC	6.21	45.38	20.53	6.24	48.43	19.47
AT/AT	3.12	6.06	23.09	3.22	4.93	26.01
TC/GA	3.22	13.57	23.01	3.26	12.17	26.10
CA/TG	3.30	−4.67	48.96	3.21	−10.37	46.24
AC/GT	3.21	−1.71	30.04	3.37	0.49	31.76
Average	3.59	9.91	29.03	3.65	8.51	30.02

−10.4°, respectively), since a positive roll of 2.2° and a slightly negative roll of about −4.7° are observed at the corresponding steps in the wild-type Cren7–DNA complex. The change in negative roll at the G1pT2 step in GR–DNA seems to have resulted from the loss of the contacts between GR and the G1:C16 pair of the DNA duplex (supplementary figure S2). On the other hand, the conformational change of Lys24 side chain, which interacts directly with the phosphate group of C6, may be responsible for the increase in negative roll at the C6pA7 step in the GR–DNA complex (Fig. 5C; supplementary figure S2).

Binding of GR also induces undertwisting of the DNA helix, which is believed to result in constraining of DNA-negative supercoils (Table 4). The base pairs at the site of intercalation show the smallest twist (~19.5°). The average twist for a single step is ~30.0° in the GR–GTGATCAC complex, which translates into ~12 bp per helical turn in bound DNA. Given the binding size of ~6 bp, binding of ~14 GR molecules would constrain one negative supercoil. For the wild-type Cren7–GTGATCAC complex, the average twist per base pair is ~29.0°, a number similar to that in the GR–GTGATCAC complex. However, due to its larger binding size (~8 bp), wild-type Cren7 is calculated to constrain one negative supercoil for ~9 proteins. Therefore, the reduced efficiency of GR in constraining DNA supercoils, as compared to that of wild-type Cren7, appears to result from a decrease in both binding size and binding affinity of the protein.

Discussion

Cren7 and Sul7d, two chromatin proteins of similar biochemical properties and tertiary structures, are both present in abundance in *Sulfolobus*. Investigation of the differences between the two proteins is of importance to the understanding of their respective roles in chromosomal organization and other DNA transactions in vivo. Structural comparison between Cren7 and Sac7d reveals that the most distinguishing feature of the two proteins in DNA-binding surface is the size of the amino acid stretch between β3 and β4 strands (Zhang et al. 2010). Notably, loop β3–β4 is highly conserved among Cren7 homologs (Fig. 1). In the present study, we show that loop β3–β4 serves a critical role in the structural integrity of and DNA binding and supercoil constraining by Cren7. Our results provide clues to the differences between Cren7 and Sac7d in biochemical properties.

All amino acid residues on loop β3–β4 that have a relatively long side chain, i.e., Pro30, Lys31, Arg33 and Lys34, contribute significantly to DNA binding by Cren7 (Table 1). Alanine substitution of Pro30 destabilized the Cren7–DNA complex, while mutation of the basic amino acid residues affected primarily the binding step of the Cren7–DNA interaction. Interestingly, the positively charged side chains of Lys31 and Lys34 are not in direct contact with the DNA in crystals (Zhang et al. 2010). So the finding that their replacement with an uncharged alanine reduced the DNA-binding affinity of the protein suggests that the positive charges on the surface of loop β3–β4 play a role in the interaction between Cren7 and DNA in solution. Further, nick-closure assays showed that the efficiencies of Cren7 mutants in constraining DNA supercoils were related proportionally to the number of basic residues on the loop (Fig. 3). The decrease in the ability of the Cren7 mutants in constraining DNA supercoils may not be attributed solely to the decrease in the affinity of these mutants for the DNA, since P30A, which showed a much reduced affinity for DNA as compared to wild-type Cren7, constrained DNA supercoils as efficiently as the wild-type protein. A possibility exists that the positively charged residues on loop β3–β4 are primarily involved in the deformation of DNA by Cren7. It is noticed that the Lys31 residue undergoes reversible methylation (Guo et al. 2008). Lysine methylation in eukaryotic histones has been shown to play a key role in the regulation of chromosomal structure and gene expression (Grant 2001; Hansen et al. 2006). A similar regulatory role may also be proposed for Lys31, which distorts the DNA backbone in the Cren7–DNA crystals through the formation of hydrogen bond to the phosphate group of the DNA.

The two highly conserved glycine residues on loop β3–β4 (Gly32 and Gly35) appear to be critical to the proper folding of Cren7. Mutation of either of the two residues led to dimerization of the protein (Fig. 4). The structural basis for the dimerization of the Cren7 mutants is unclear. It has been reported that mutation of a glycine residue (Gly55) promotes the conversion of the IgG-binding domain of protein L from a monomeric form into a dimeric form at a protein concentration of above 30 μM (Kuhlman et al. 2001). The 62-residue domain consists of an α-helix packed on a four-stranded β-sheet formed by two β-hairpins. Alanine substitution of Gly55, located in the C-terminal hairpin, straightens the four-residue turn in this hairpin, allowing the C-terminal β strand to extend out and insert into a partner protomer to form a domain-swapped dimer. Conceivably, the mutant Cren7 proteins may form dimers in a similar manner to the IgG-binding domain of protein L. It is possible that, once the glycine residue is mutated, loop β3–β4 becomes less flexible, and the straightened loop permits the β4 and β5 strands to flip out. The two β strands then form an intact β barrel with the β1, β2 and β3 strands from another Cren7 mutant protomer. Dimerization of the mutant Cren7 protein would presumably be favored in vivo, given the millimolar levels of the protein present in *Sulfolobus*. Therefore, the two glycine residues appear to

be essential in maintaining the flexibility of loop β3–β4 and the monomeric form of Cren7 in vivo.

The presence of a much longer loop between β3 and β4 strands in Cren7 than that in Sul7d raises a question about the role of the size of the loop on the properties of the proteins. Deletion of five residues in the middle of loop β3–β4 resulted in drastic loss in both DNA-binding and supercoil-constraining abilities of Cren7. The loss of function of the shortened loop was not effectively compensated even when the loop was a small two-residue hinge (i.e., Gly29–Arg30 for GR or Gly29–Lys30 for GK), which mimicked that in Sac7d (Figs. 2, 3). Comparison of the crystal structures of wild-type Cren7 and GR in complex with dsDNA reveals the structural basis for the possible functional differences between the long loop and the small hinge. First, GR occupies a binding site (~6 bp) smaller than that of wild-type Cren7 (~8 bp). Second, no base pair contacts by GR were observed in the region from G1-C16 to G3-C14, and thus the DNA strand appears much smoother than that bound by the wild-type protein. Third, DNA is less bent in the GR–DNA complex (~59.6°) than in the wild-type Cren7–DNA complex (~69.4°). Given the potential effect of DNA bending on DNA compaction, possession of the long loop is important for Cren7 to play a role in chromosomal DNA organization. Fourth, GR but not Cren7 narrows the major groove at the intercalation site. In this regard, GR shows greater resemblance to Sul7d than to Cren7. In fact, as compared to Cren7, both GR and Sul7d possess smaller DNA-binding sizes, and bend DNA without severely distorting the backbone. An exception is that the DNA curvature induced by GR is smaller than that by Sul7d. We infer from these results that the size of loop β3–β4 may influence the width of the major groove of dsDNA bound by the chromatin protein and, therefore, the chromosomal accessibility of other DNA-binding proteins that function in various DNA-associated activities. The distinct difference in the impact between Cren7 and Sul7d on the major groove of DNA is consistent with non-redundant functions of the two chromatin proteins in DNA packaging and transactions.

Acknowledgments We are grateful to Dr. Peng Cao of the Institute of Biophysics, Chinese Academy of Sciences for her help in structure refinement. This work was supported by grant 31130003 to Li Huang and grants 30900023 and 31270124 to Zhenfeng Zhang from the National Natural Science Foundation of China.

References

Arnott S (1999) Oxford Handbook of Nucleic Acid Structure. Oxford University Press, Oxford

Bell SD, Botting CH, Wardleworth BN, Jackson SP, White MF (2002) The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. Science 296:148–151

Collaborative Computational Project N (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50:760–763

Cubonova Lu, Sandman K, Hallam SJ, Delong EF, Reeve JN (2005) Histones in crenarchaea. J Bacteriol 187:5482–5485

Dame RT (2005) The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. Mol Microbiol 56:858–870

Driessen RPC et al (2013) Crenarchaeal chromatin proteins Cren7 and Sul7 compact DNA by inducing rigid bends. Nucleic Acids Res 41:196–205

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126–2132

Feng Y, Yao H, Wang J (2010) Crystal structure of the crenarchaeal conserved chromatin protein Cren7 and double-stranded DNA complex. Protein Sci 19:1253–1257

Gao YG et al (1998) The crystal structure of the hyperthermophile chromosomal protein Sso7d bound to DNA. Nat Struct Biol 5:782–786

Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15:305–308

Grant PA (2001) A tale of histone modifications. Genome Biol 2: REVIEWS0003

Grote M, Dijk J, Reinhardt R (1986) Ribosomal and DNA-Binding Proteins of the Thermoacidophilic Archaebacterium *Sulfolobus Acidocaldarius*. Biochim Biophys Acta 873:405–413. doi:10.1016/0167-4838(86)90090-7

Guo R, Xue H, Huang L (2003) Ssh10b, a conserved thermophilic archaeal protein, binds RNA in vivo. Mol Microbiol 50:1605–1615

Guo L, Feng Y, Zhang Z, Yao H, Luo Y, Wang J, Huang L (2008) Biochemical and structural characterization of Cren7, a novel chromatin protein conserved among Crenarchaea. Nucleic Acids Res 36:1129–1137

Hansen JC, Lu X, Ross ED, Woody RW (2006) Intrinsic protein disorder, amino acid composition, and histone terminal domains. J Biol Chem 281:1853–1856

Kuhlman B, O'Neill JW, Kim DE, Zhang KY, Baker D (2001) Conversion of monomeric protein L to an obligate dimer by computational protein design. Proc Natl Acad Sci USA 98:10687–10691

Laskowski RA, Macarthur MW, Moss DS, Thornton JM (1993) Procheck—a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291. doi:10.1107/S0021889892009944

Lu X-J, Olson WK (2008) 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. Nat Protoc 3:1213–1227

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature 389:251–260

Luijsterburg MS, White MF, van Driel R, Dame RT (2008) The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes. Crit Rev Biochem Mol Biol 43:393–418

Luo X, Schwarz-Linek U, Botting CH, Hensel R, Siebers B, White MF (2007) CC1, a novel crenarchaeal DNA binding protein. J Bacteriol 189:403–409

Mai VQ, Chen X, Hong R, Huang L (1998) Small abundant DNA binding proteins from the thermoacidophilic archaeon *Sulfolobus shibatae* constrain negative DNA supercoils. J Bacteriol 180:2560–2563

Marsh VL, Peak-Chew SY, Bell SD (2005) Sir2 and the acetyltransferase, Pat, regulate the archaeal chromatin protein, Alba. J Biol Chem 280:21122–21128

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. J Appl Crystallogr 40:658–674. doi:10.1107/S0021889807021206

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53:240–255

Napoli A, Zivanovic Y, Bocs C, Buhler C, Rossi M, Forterre P, Ciaramella M (2002) DNA bending, compaction and negative supercoiling by the architectural protein Sso7d of *Sulfolobus solfataricus*. Nucleic Acids Res 30:2656–2662

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. Method Enzymol 276:307–326. doi:10.1016/S0076-6879(97)76066-X

Reeve JN (2003) Archaeal chromatin and transcription. Mol Microbiol 48:587–598

Robinson H, Gao YG, McCrary BS, Edmondson SP, Shriver JW, Wang AH (1998) The hyperthermophile chromosomal protein Sac7d sharply kinks DNA. Nature 392:202–205

Sandman K, Reeve JN (2005) Archaeal chromatin proteins: different structures but common function? Curr Opin Microbiol 8:656–661

Wu LJ (2004) Structure and segregation of the bacterial nucleoid. Curr Opin Genet Dev 14:126–132

Zhang Z, Gong Y, Guo L, Jiang T, Huang L (2010) Structural insights into the interaction of the crenarchaeal chromatin protein Cren7 with DNA. Mol Microbiol 76:749–759

Zhang Z, Guo L, Huang L (2012) Archaeal chromatin proteins. Sci China Life Sci 55:377–385